**The Thesis Committee for Jacob John Herman**
**Certifies that this is the approved version of the following thesis:**

# Morphometrics of the Dwarf Honey Bee *Apis florea* show Biogeographic Differentiation across India

**APPROVED BY**
**SUPERVISING COMMITTEE:**

**Supervisor:**

Ulrich Mueller

Craig Linder

# Morphometrics of the Dwarf Honey Bee *Apis florea* show Biogeographic Differentiation across India

**by**

**Jacob John Herman**

**Thesis**

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

**Master of Arts**

**The University of Texas at Austin**

**December 2017**

# Acknowledgements

This study was funded by the National Center for Biological Sciences institutional funding, Texas Ecolab, and EEB startup grant. I would like to thank Hanumantha Raju and the many beekeepers that helped in locating and collecting from wild colonies. I thank Emily Wissel for help with statistical analysis. I thank Ulrich Mueller for suggestions on improving the manuscript and lab space for analysis.

# Abstract

## Morphometrics of the Dwarf Honey Bee *Apis florea* show Biogeographic Differentiation across India

Jacob John Herman, M.A.

The University of Texas at Austin, 2017

Supervisor: Ulrich Mueller

The Asian dwarf honey bee (*Apis florea*) is a relatively small honey bee, nests in the open with single combs attached to tree branches, and inhabits areas uninhabitable to other *Apis* species. *A. florea* is one of few honey bees in the genus to have remained unmanaged by beekeepers across Asia. Because *A. florea* has not been bred for specific traits or transported intentionally across continents like managed *Apis* species, populations of *A. florea* should offer insight into natural adaptations of honey bee populations to diverse climates. We use morphometrics to examine which environmental factors correlate with morphological differences between populations of *A. florea* surveyed across India. The surveyed populations show a trend of increased wing size going from the equator to the north. The populations also vary in Cubital Index, a wing venation measurement often associated with subspecies differentiation, and this variation is correlated with minimum temperature of the coldest month. Taken together, these findings show that *A. florea* differs morphologically across a temperature gradient in

India and support future work towards understanding biogeographic patterns in this understudied species of honey bee.

# Table of Contents

# List of Tables

## List of Figures

## 1. Introduction

Understanding natural variation in honey bee populations is important to address and manage honey bee health concerns (Meixner 2010) because it allows us to see how honey bees can resist many harsh environments, past, present and future (Seeley 2014). Unfortunately, honey bee colonies are very difficult to study in natural conditions free of human influence because many colonies found in nature are either swarms absconded from beekeepers, or have mated with drones from managed honey bee genotypes (Tarpy et al. 2015). Because of these apicultural influences, honey bee population structure is greatly influenced by human-mediated bee transport and gene flow.

*A. florea* is a better study species than other honey bees, in that it has been largely unmanaged, and has relatively short foraging and mating distances (Dyer and Seeley 1991). The distribution of the dwarf honey bee, *A. florea*, has previously been mapped throughout Southeast Asia, with its range encompassing much of Asia's most diverse climates, from as far west as Eastern Africa, and as far east as Eastern China (Ruttner 1988, Hepburn et al. 2005). The first step towards understanding these natural populations is identifying variable morphological indicators that correlate with environmental variation. Fortunately, within *Apis*, there are several measurements of wing venation that largely reflect population divergences (Tofilski 2008, Oleksa and Tofilski 2015), and can be measured relatively easily using images of the forewings and available morphology measuring software.

This study aimed to improve understanding of environmental factors that influence differentiation of *A. florea* within India, and to identify candidate populations for future study of local adaptations. I expected that, compared to human-managed *Apis* species, natural geographic differentiation and local adaptation in *Apis florea* would be preserved because of (a) significant variation in temperature, precipitation, and other environmental variables across India, (b) the small mating and foraging ranges relative to other *Apis* species, and (c) the lack of management and transport by humans.

## 2. Methods

*Apis florea* colonies are typically found exposed on tree branches or other structures with the colony consisting of a single comb (Figure 1). The comb is roughly organized into honey stores towards the top around a single branch, and pollen and larvae inside cells spanning the rest of the comb. This comb construction helps facilitate the collection of bees of similar age, because most foragers will fly from the top platform of the comb (crown) when the colony is disturbed (Hepburn et al. 2014).

### Bee Collection

Bees were collected from individual colonies from February to June 2016 from 55 locations throughout India (Figure 2, Table 1). Collections involved taking 10-20 bees non-destructively from each colony by sweeping an adhesive strip over the crown of a colony and collecting bees that adhered to the adhesive strip upon upward flight. This method of collection allowed for minimal disturbance of colonies, with no colonies absconding after collection. Bees were preserved in ethanol for transport and later stored in a 4 °C freezer.

### Environmental Data Collection

GPS locations were taken in the field at each colony site using a Garmin 64s GPS, which has an average accuracy of 3 meters (Garmin International). These coordinates were then used to extract local environmental factors from WorldClim (Hijmans 2004), including

precipitation, annual temperature, elevation, and available bioclimatic variables (Supplementary Table 1).

### Measuring Techniques

The left forewing of each bee was mounted individually on a microscope slide. Mounted wings were photographed using an Olympus TG-4 microscope (Figure 3). A program that measures 2D shapes, tpsDig2.16 (Rohlf 2005), with set landmarks of specific sites (in this case wing venation terminations) was then used to generate TPS files for each sample. Landmarks chosen were those of joining vein intersections, following previous studies (Tofilski 2015), as indicated in Figure 3. Distances between major vein-intersections were then found within tpsDig2.16 in order to measure vein lengths (Supplementary Table 2). Cubital Index, which has been used to cluster populations in other *Apis* species (Ruttner 1978) was calculated by dividing length 1-2 by length 2-4 (Figure 3). Additionally, a random subset of samples was blindly re-measured to evaluate accuracy in measuring, which resulted in minimal error of approximately 0.02%. (Figure 4).

## Statistical Analysis

### Principal Component Analyses of Morphology and Environmental Variables

For each sample, all vein lengths, wing lengths, and cubital indices were recorded (Figure 3). To account for colony variation, 5-10 samples collected from the same colony were

averaged to generate a single score per individual colony. A Principle Component Analysis (PCA) was then conducted to test for clustering of samples by morphologies within R (R Core Team 2013). Environmental data exported from WorldClim climate grid data were imported, and likewise analyzed in a PCA to test whether environmental data at the collection sites clustered together.

### ANOVA of Cubital Index and Wing Length

Splitting samples into Northern India and Southern India, an ANOVA was conducted for Cubital Index from northern and southern samples, and then repeated for wing length. In a secondary analysis northern samples were split further into northwest and northeast sub-samples.

### Regression Analysis of Cubital Index and Wing Length

To evaluate the predictive power of each environmental variable, a backwards selection linear regression was performed for both the prediction of Cubital Index, and separately for the prediction of wing length. This analysis was done in R with the LM function. This analysis started by including all environmental variables, then removed the least significant variables one by one until all remaining environmental variables were significant. After this process, further testing for collinearity between environmental variables was performed with the Variant Inflation Factors (VIF) function in R.

## 3. Results

### Morphology PCA:

The morphology PCA shows no clear clustering when using PC1 and PC2, with most of the diversity in samples falling along PC1 (Figure 5). The arrow plot of this PCA (Figure 6) shows that Cubital Index explains variance largely in the opposite direction of all other measurements scoring along PC1. PC1 reflects ~42.19% of total variance and PC2 ~8.16% (Figure 7).

### Environmental PCA:

No clear clustering emerged in a PCA of all environmental variables measured for the sampled locations (Figure 8). The arrow plot of this PCA (Figure 9) shows that Elevation is contributing largely to PC2 variance, with no other clear separation of factors. PC1 explains ~52.51% of total variance, and PC2 explains ~21.32% of total variance (Figure 10).

### ANOVA Wing Length and Cubital Index:

ANOVAs for wing length and Cubital Index were conducted comparing northern and southern samples. This yielded statistical significance for wing length across the samples (P=.0.0533), and no statistical significance for Cubital Index (P=0.8034). When northern samples were further split into northeast and northwest populations, wing length became statistically non-significant (P=0.09793), and Cubital Index remained statistically non-

significant (P=0.5204). However, the northern split resulted in significantly smaller sample size for the northeast cluster, reducing statistical power.

**Backwards Stepwise Regression Analysis:**

*Wing Length*: After eliminating non-signifcant predictors of wing length, only environmental variables Latitude (P=0.00007.03), Elevation (P=0.07608), and Mean Temperature 4 (P=0.0009387) (Table 2) remained significant. To minimize the effects of collinearity, all significant variables fall below a cutoff of 4 for VIF (Table 2).

*Cubital Index:* Using the same elimination method, the only variables contributing significantly to cubital index were Temperature of the Coldest Month (bio_6) (P=0.004935), Elevation (P=0.362835), and Latitude (P=0.184326) (Table 2). Accounting for collinearity, only Elevation (VIF=2.47555) remains after a cutoff of VIF=4.

## 4. Discussion

My survey of *Apis florea* across India showed that latitude, temperature, and elevation largely predict morphological variation in *Apis florea* (Table 2). However, the principal component analysis (PCA) of wing morphology showed no clear clustering of populations by morphologies, even when accounting for latitude (Figure 5). This is likely because Cubital Index contributes to variance differently than all other measurements, as seen in the arrow matrix (Figure 6). From the backwards stepwise regression (Table 2), where variables with collinearity were removed, followed by removal of environmental variables that were not significant, Cubital Index correlated with elevation, latitude, and minimum temperature of the coldest month (bio_6), while wing length was predicted by elevation, latitude, and overall mean temperature. Variation in wing length is particularly interesting, because wing size is indicative of the overall size of the bees. Like other studies (Hepburn et al. 2005), I found smaller bees in the southernmost lowlands of India near the equator, with bee size increasing as latitude increases. To date, no studies have addressed whether these morphological differences in *A. florea* are due to plastic developmental responses to environmental differences, or whether bees in the different locations are genetically different to explain these morphological differences.

The results of the PCA of morphology is consistent with many previous studies of the genus *Apis* (Ruttner 1986, Zhu et al. 2017, Rinderer et al. 1989), in that Cubital Index is the most variable parameter of the wing venation measurements across populations. A

previous study of *A. florea* found little morphological differentiation across the entire range from Africa to Eastern China (Hepburn et al. 2005), but our analysis of Indian populations shows there exists significant morphological variation even at the smaller scale of the country of India.

Of the variation that I observed, Cubital Index seems to vary largely with minimum temperature. Unlike cavity nesters, where heat generated in the hive is conserved in enclosed spaces, *A. florea* is an open-nesting bee (Figure 1), and this open nesting may lead to differences in individual physiological thermoregulation in colder climates, rather than behavioral thermoregulation as in cavity nesters. Future studies could examine if variation present in *A. florea*'s coldest ranges is a result of adaptation to those regions, and if the adaptations are at an individual or whole colony level.

This study is first to tie *A. florea* population variation to known variation found through modern GIS techniques, and differs from previous studies in the greater sampling of regional diversity, and sampling directly from nest sites. Many previous studies have collected foraging *A. florea* either in the field, or not reported how the bees were collected. I hope that future studies will follow our sampling scheme of collecting multiple bees per colony for any molecular work addressing within-colony and population-genetic variation. While collecting multiple samples from the same colony is beneficial, it will also be important that bees from multiple colonies are collected in a given area. This was done for a few locations in this study, but the study would likely have benefited from repeated sampling at more locations, particularly at intermediate

latitudes across India, and north-central India. Additionally, while morphology is very informative to understanding population differentiation, stronger evidence would come from analysis of genetic variation at neutral genetic loci and the biogeographic distribution of this variation across the Indian range. Without molecular support, any Cubital Index and other wing morphology changes, could be a result of plastic developmental responses and not underlying genetic differences. I hope to address this in future molecular work analyzing population-genetic structure.

This study lends itself to several follow-up studies. To address local adaptation, an interesting future study would be a transfer experiment or a common-garden experiment. For example, relocating queens or colonies from one climate extreme, and transporting them to regions of the other climate extreme, would allow researchers to observe how the bees develop and behave in the changes in conditions. Common-garden experiments have been attempted a few times with other bee species (Field 2012), but have never been tried with *A. florea*. *Apis florea* offers comparative insight into the evolution of honeybees, and may hold key information about *Apis* adaptation, because it is able to live in some of the harshest climates of all *Apis* species (Hepburn et al. 2005).

In summary, I have shown that combining traditional morphological measurements for *Apis* with world-climate data can help predict morphological variation in natural *Apis* populations. This is the first study to tie *A. florea* population variation to environmental variation using GIS information. Future studies that quantify physiology and genomic differentiation should help reveal how *A. florea* is able to live in the extreme

climates it occupies, and if selection in these diverse environments enabled local adaptation in these populations. In addition, understanding of *A. florea's* adaptations to extreme climates should also increase understanding how other species in the genus can potentially adapt to extreme climates, present and future.

**Figure 1.** *Apis florea* nest with bees removed to show comb structure. The comb is suspended from a branch near the top. The portion of the comb wrapped around the branch is called the crown. Photo by Ravi Kumar Boyapati.

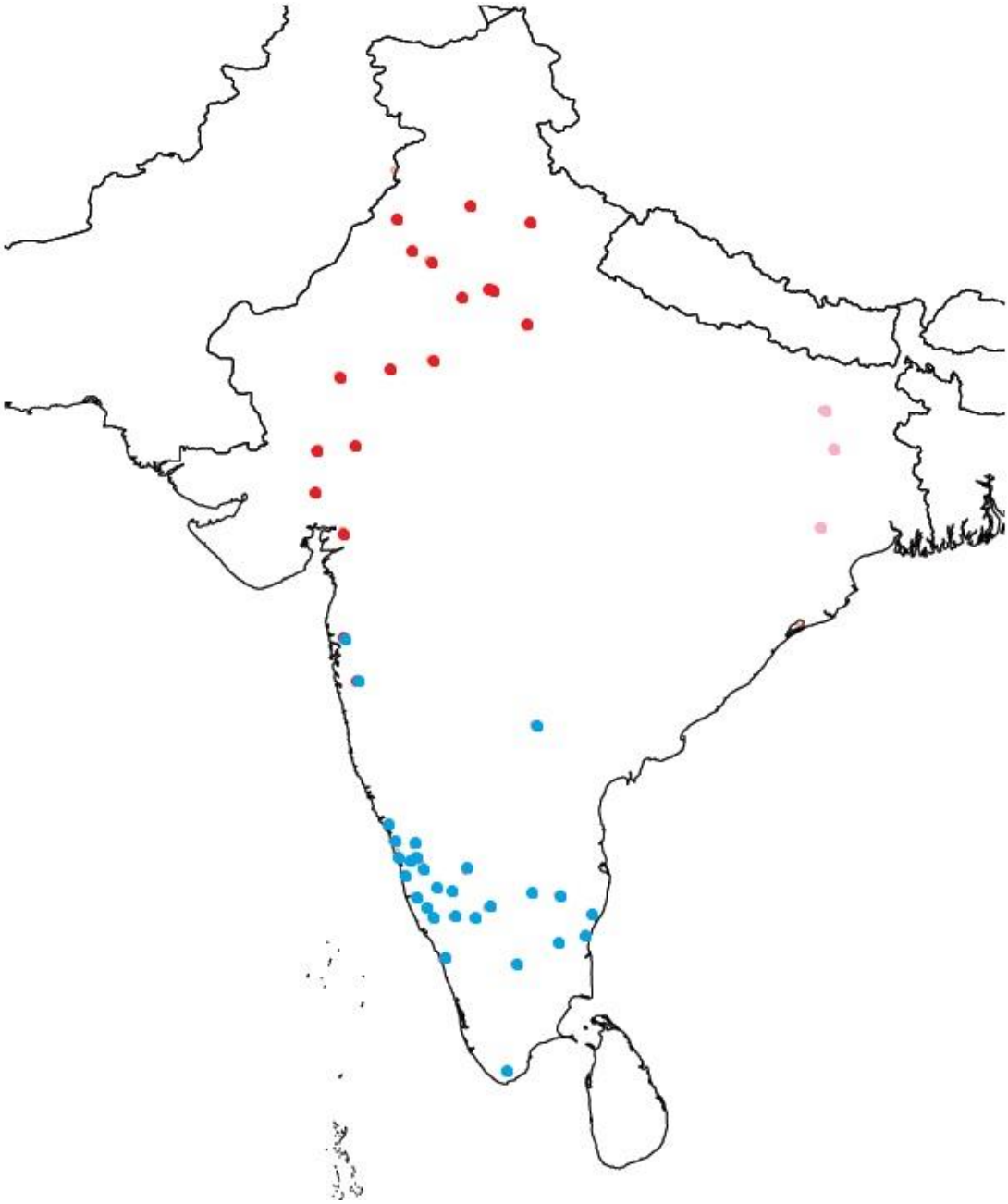

**Figure 2.** *Apis florea* collection sites within India. Northern samples are red and pink; southern samples are blue. Northern samples are further divided into northwest (red) and northeast (pink).

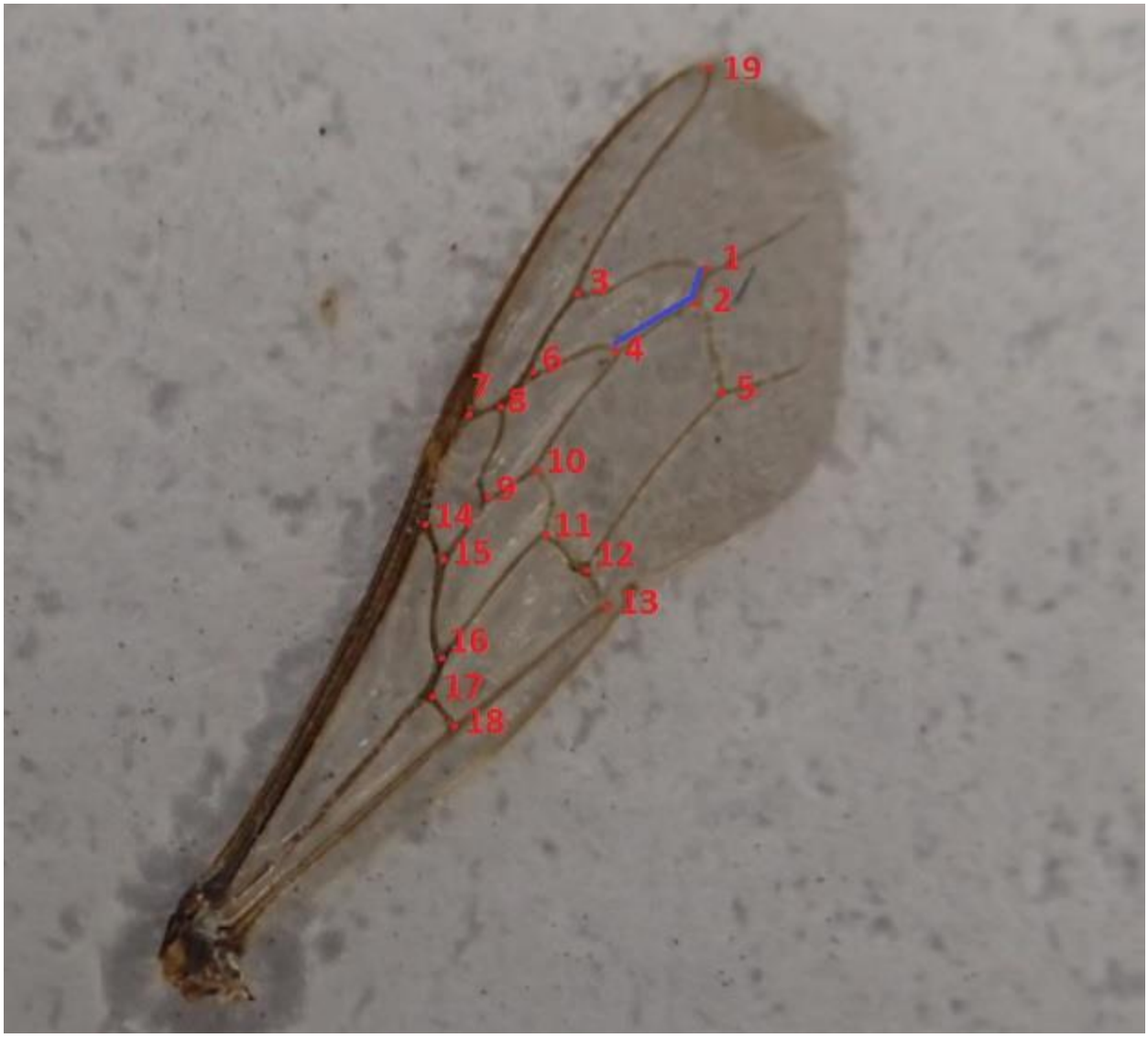


**Figure 3.** *A. florea* wing with the numbering of wing venation junctions used in the morphometric analyses. The Cubital Index is calculated by dividing the length between points 1-2 by the length between points 2-4, shown in blue lines here.

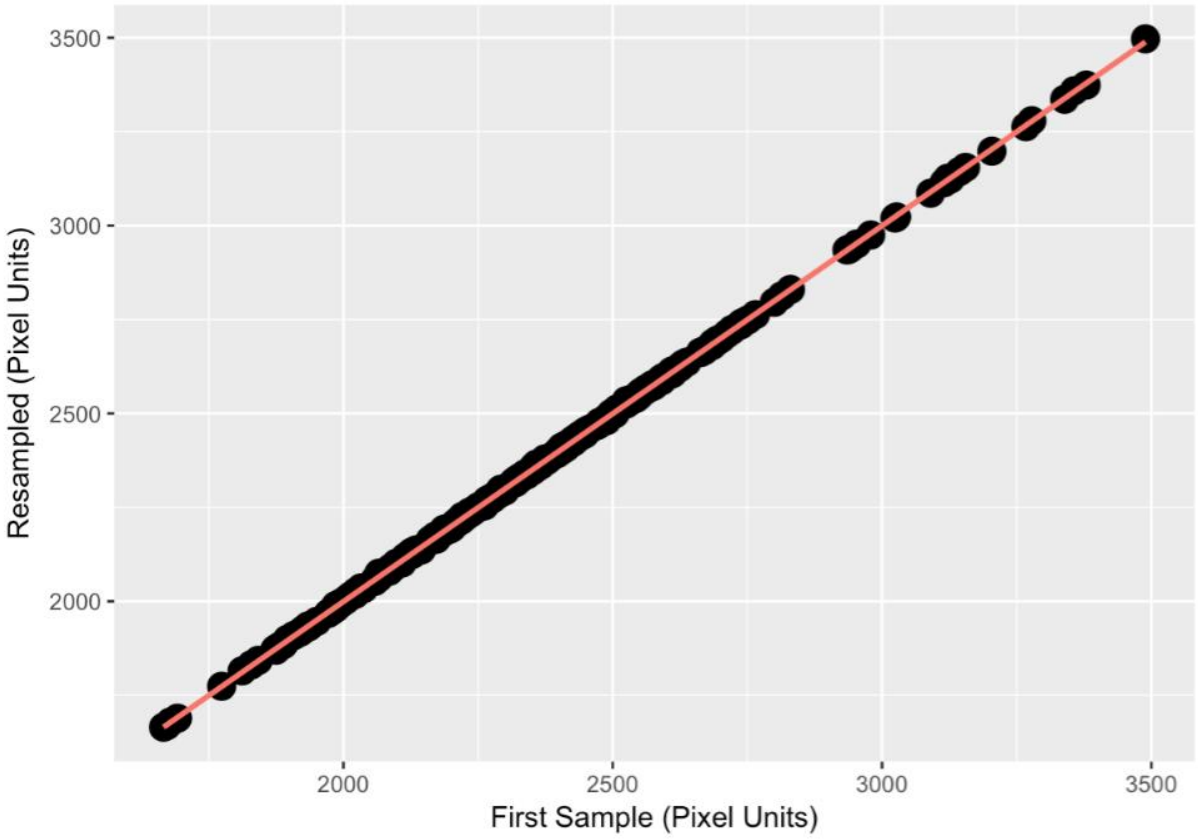


**Figure 4.** Wing re-measurements plotted for assessment of measurement error. Units of both axes represent coordinates within tpsDig2.16, recorded in pixel units. Re-measurements include 267 coordinates (pictured), with percent differences averaging 0.02% overall. A best fit line is given in red.

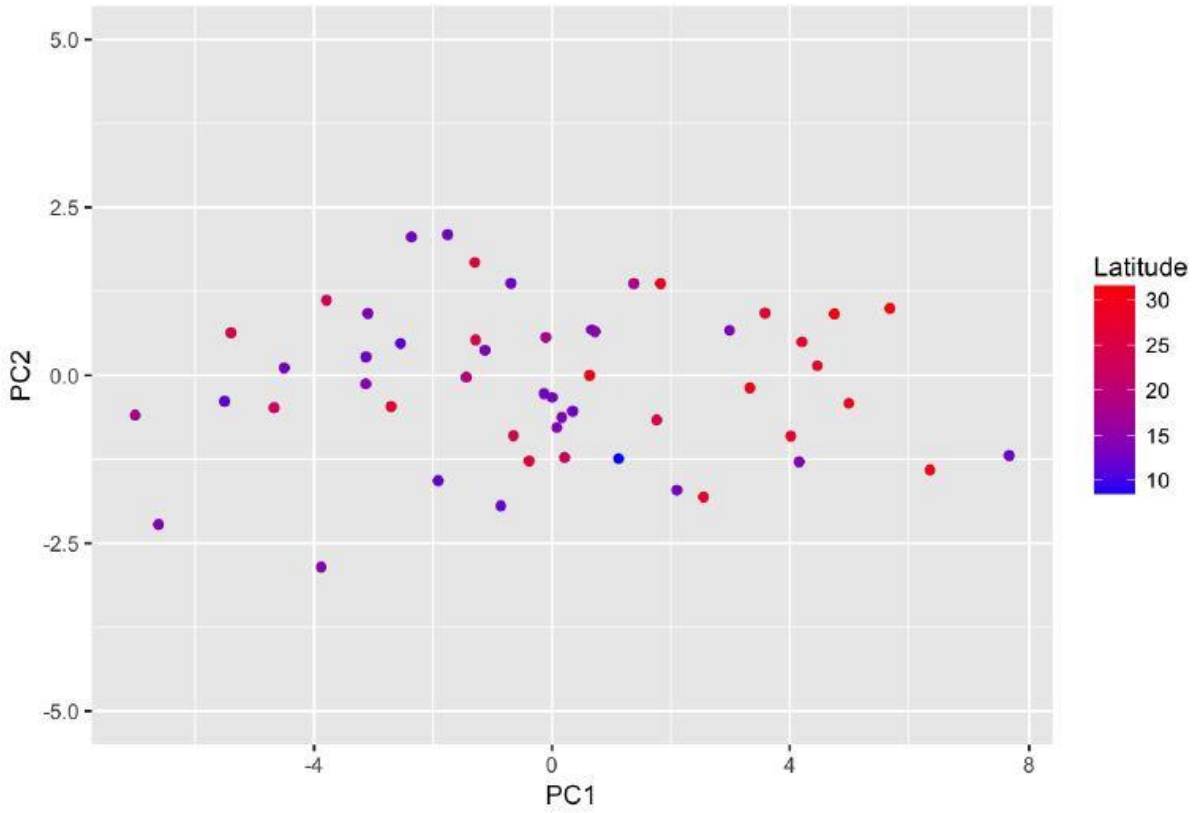


**Figure 5.** Principal component analysis of wing venation measurements labeled by corresponding latitude. This figure was generated in R to emphasize the variation in the data set and to look for patterns of in wing measurements against latitudes. There is no clear clustering by latitude.

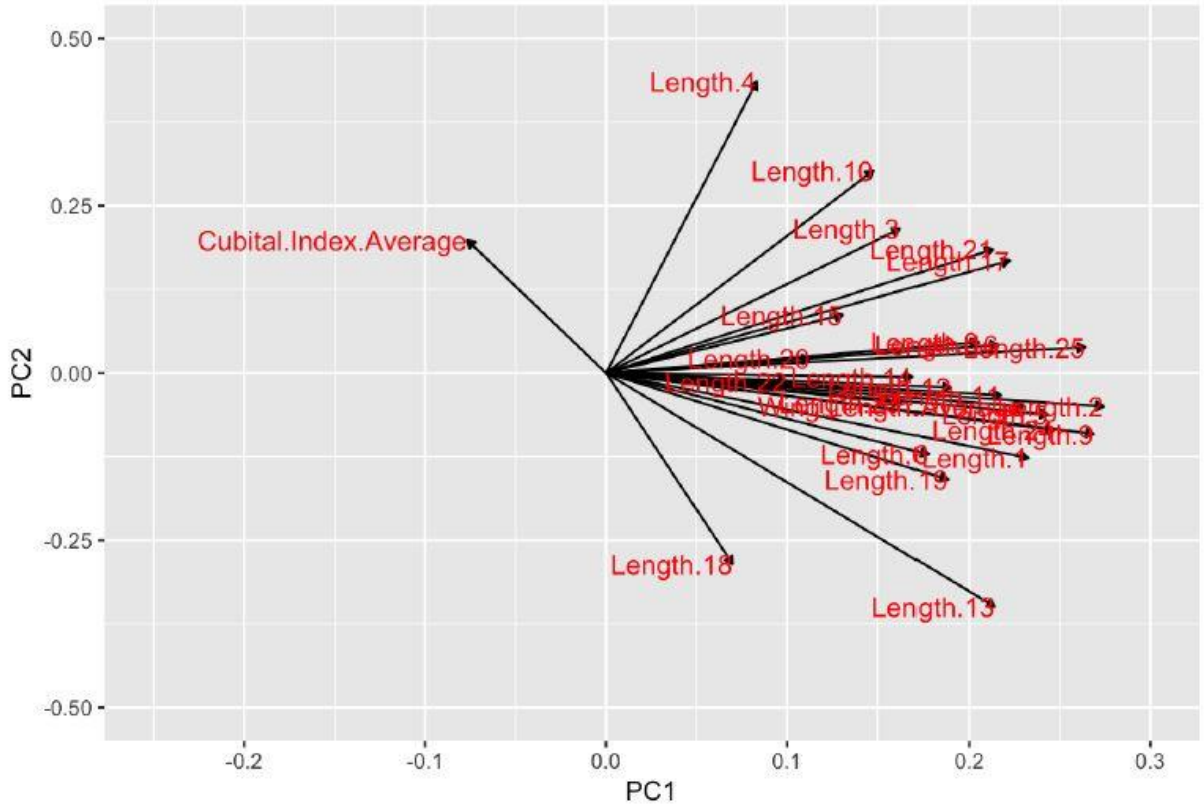


**Figure 6**. Arrow plots display how each variable independently contributes to variance in a principal component analysis (PCA). This arrow plot complements the PCA in Figure 5, and shows that Cubital Index Average contributes to variance differentially from all other measurements. Points on the left side of the PCA tended to have larger Cubital Index averages, while points on the right side tended to have higher scores on all other measurements. Because there is no distinct clustering, the strength of this relationship is weak.

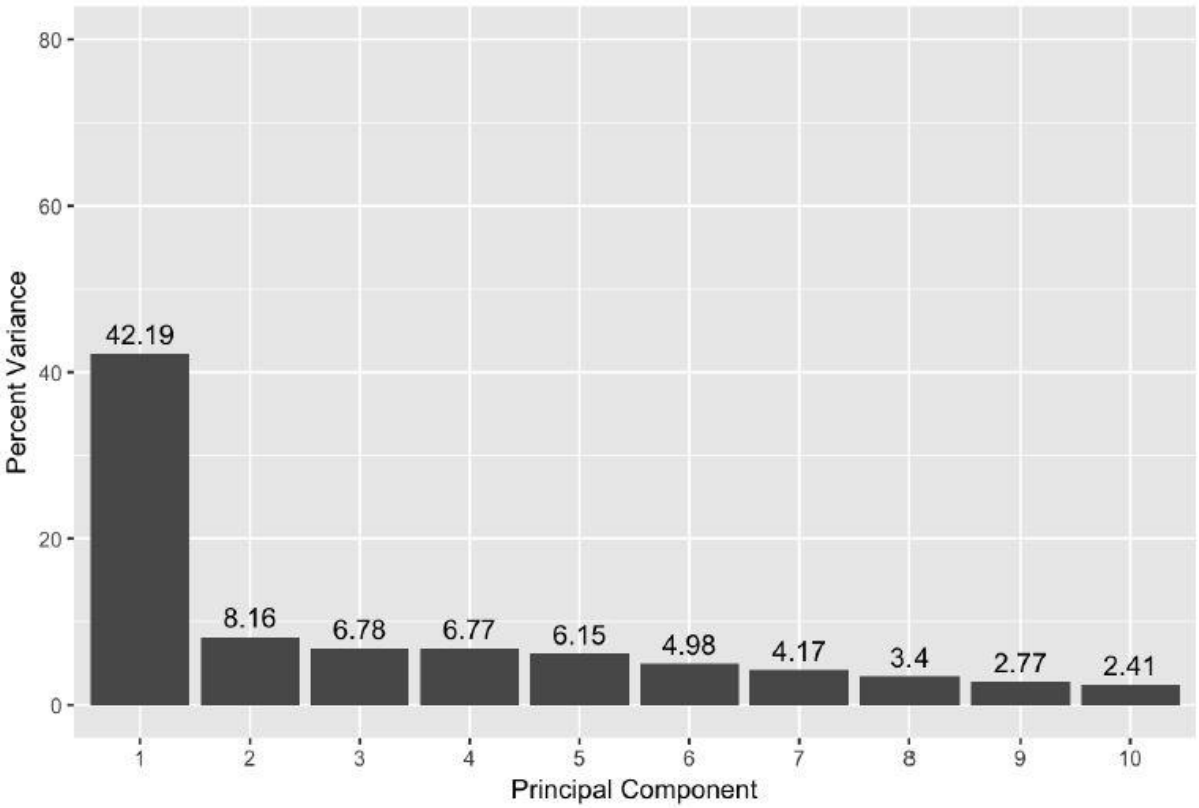


**Figure 7**. Percentage of variance explained by each principal component (PC) in the PCA in Figure 5. The PCA in Figure 5 represents 50.35% of the total variance observed in the data set, with the top 10 PCs explaining most of the variance in the data.

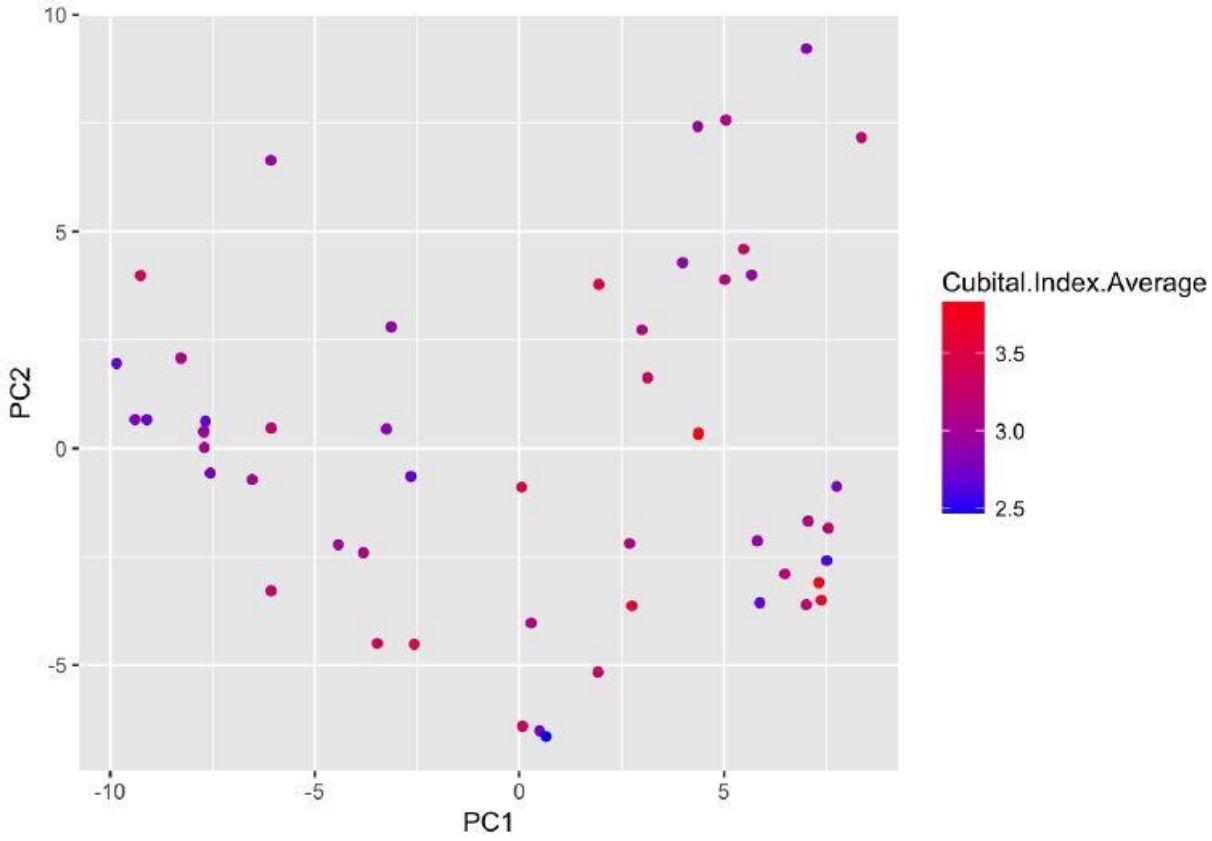


**Figure 8.** This figure examines how environmental variables explain Cubital Index average across populations. As in Figure 5, there is no clear clustering by latitude.

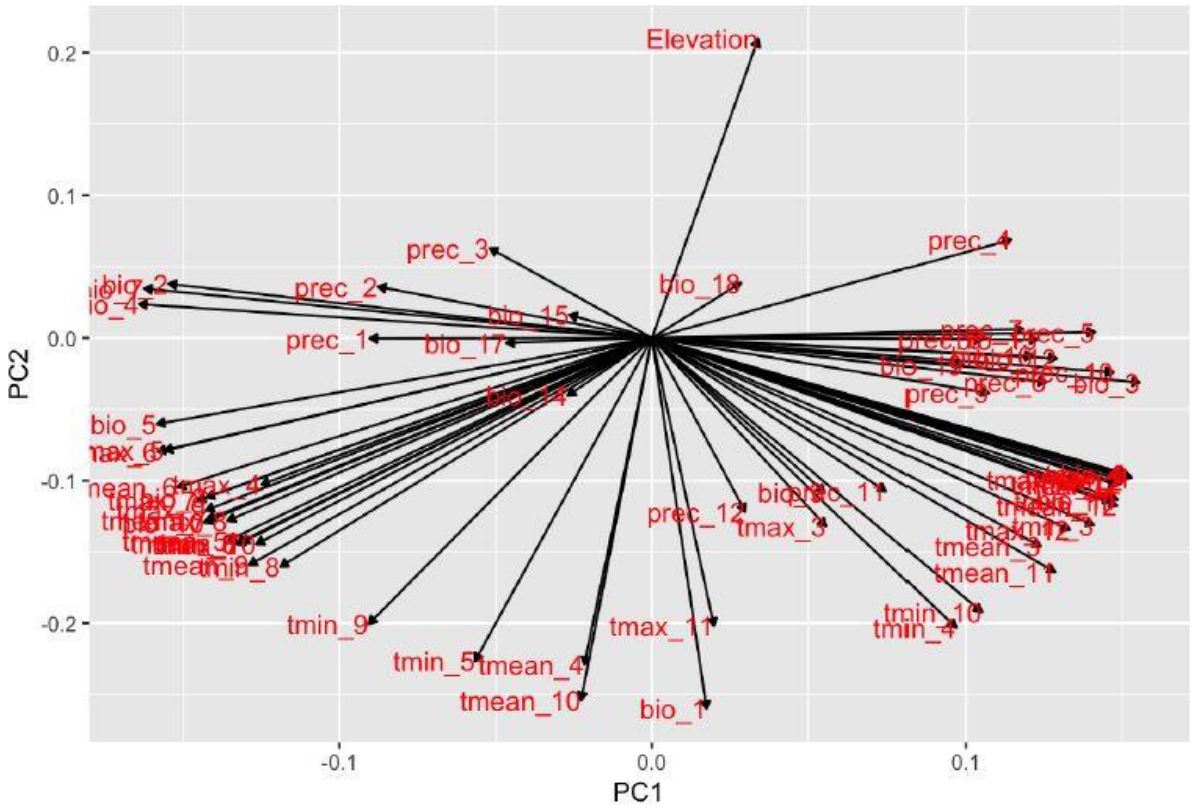


**Figure 9.** Arrow plot displays how each variable contributes to variance seen in the PCA of environmental variables in Figure 8. Here, elevation seems to contribute to variance distinctly from the rest of the variables, indicating that higher elevation is associated with data points having a larger Y value in the PCA of environmental variables in Figure 8. As in Figure 6, this effect may be subtle, as distinct clusters are not observed in the PCA.

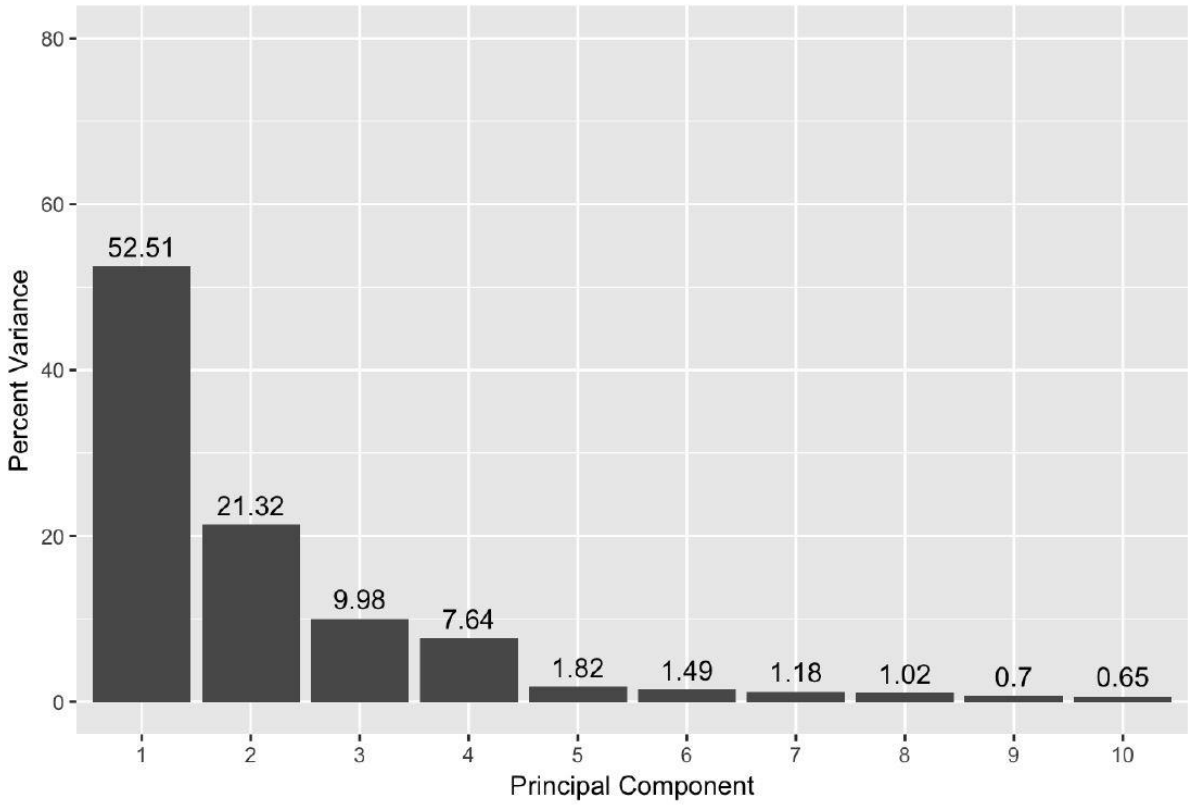


**Figure 10.** Percentage of variance explained by each PC in the PCA in Figure 8. The PCA in Figure 8 represents 73.83% of the total variance seen in the data set, with the top 10 PCs explaining most of the variance in the data.

**Table 1.** Collection locations of sampled colonies. Longitude and latitude are given in degrees east and north, respectively. Location names list the nearest city to each collection site.

| Collection Location | Longitude | Latitude |
|---|---|---|
| Calicut | 75.802 | 11.255 |
| Keeriparai | 77.512 | 8.414 |
| Vadodara | 73.110 | 22.187 |
| Amedbad | 72.392 | 23.101 |
| Udaipur | 73.415 | 24.343 |
| Base Abu Road | 72.474 | 24.292 |
| Jodpur | 73.001 | 26.161 |
| Ashnesi | 74.384 | 26.280 |
| Jaipur | 75.491 | 26.555 |
| Riwari | 76.367 | 28.121 |
| Hisar | 75.428 | 29.085 |
| Mandya | 77.013 | 12.573 |
| Mysore | 76.638 | 12.308 |
| Pirayapattan | 76.079 | 12.342 |
| Bhagamandala | 75.510 | 12.370 |
| Sullia | 75.370 | 12.580 |
| Puttur | 75.132 | 12.819 |
| Shambur | 75.069 | 12.870 |
| Brahmavara | 74.748 | 13.440 |
| Kollur (Mudar) | 74.870 | 13.790 |
| Bhatkal | 74.555 | 14.002 |
| Honnovar | 74.450 | 14.280 |
| Ankola | 74.304 | 14.626 |
| Ankola | 74.297 | 14.669 |
| Sagar | 75.027 | 14.163 |
| Hasanagara | 75.065 | 13.912 |
| Thirthahalli | 75.245 | 13.691 |
| Mudigere | 75.640 | 13.135 |
| Hassan | 76.070 | 13.010 |
| Sirsa | 75.010 | 29.318 |
| Bhatinda | 74.552 | 30.127 |
| Amritsar | 74.521 | 31.378 |
| Pinjone 1 | 76.546 | 30.470 |
| Pinjone 2 | 76.545 | 30.471 |
| Rishikesh | 78.179 | 30.067 |

**Table 1.** Continued

| | | |
|---|---|---|
| Delhi 1 | 77.142 | 28.394 |
| Delhi 2 | 77.056 | 28.385 |
| Delhi 3 | 77.113 | 28.310 |
| Hathros | 78.083 | 27.377 |
| Mumbai | 72.960 | 19.369 |
| Pune 1 | 73.484 | 18.324 |
| Pune 2 | 73.485 | 18.327 |
| Near Bangalore | 76.444 | 13.504 |
| Hyderbad | 78.288 | 17.248 |
| Tiring | 86.048 | 22.314 |
| Jasidhi | 86.387 | 24.310 |
| Begusaai | 86.082 | 25.254 |
| Bhangadpet | 78.226 | 12.991 |
| Palikonda | 78.935 | 12.902 |
| Memaruvvathuv | 79.826 | 12.415 |
| Puducherry | 79.675 | 11.919 |
| Bangaram | 78.911 | 11.672 |
| Sivagiri | 77.762 | 11.152 |

**Table 2.** Summary of backwards stepwise regression of Cubital Index and wing length average on WorldClim variables. Asterisks denote significance. WorldClim variable bio_6 is statistically significant in predicting Cubital Index. Latitude and mean temperature are statistically significant in predicting wing length.

| Response: | Cubital Index. | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Df | Sum Square | Mean Square | F-value | Pr(>F) | Significance | VIF | Effect Size (Cohen's $F^2$) |
| Elevation | 1 | 0.0788 | 0.07877 | 0.8437 | 0.362835 | | 2.47555 | 0.0048108 |
| Latitude | 1 | 0.1693 | 0.16927 | 1.8132 | 0.184326 | | 22.30839 | 0.0845746 |
| bio_6 | 1 | 0.8094 | 0.80942 | 8.67 | 0.004935 | ** | 21.5626 | 0.1461076 |
| Residuals | 49 | 4.5745 | 0.09336 | | | | | |

| Response: | Wing Length | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Df | Sum Square | Mean Square | F-value | Pr(>F) | Significance | VIF | Effect Size (Cohen's $F^2$) |
| Elevation | 1 | 0.07219 | 0.07219 | 3.2841 | 0.0760847 | | 1.643959 | 0.0432811 |
| Latitude | 1 | 0.41471 | 0.41471 | 18.8667 | 7.03E-05 | *** | 1.0716 | 0.1986826 |
| tmean_4 | 1 | 0.27259 | 0.27259 | 12.4012 | 0.0009387 | *** | 1.583048 | 0.2916305 |
| Residuals | 49 | 1.07708 | 0.02198 | | | | | |

| Significance codes: | <0.001 '***' | 0.001 '**' | 0.01 '*' | 0.05'.' | 0.1 ' ' 1 |
|---|---|---|---|---|---|

**References**

Dyer, F. C., & Seeley, T. D. (1991). Dance dialects and foraging range in three Asian honey bee species. *Behavioral Ecology and Sociobiology*, *28*(4), 227-233.

Field, J., Paxton, R., Soro, A., Craze, P., & Bridge, C. (2012). Body size, demography and foraging in a socially plastic sweat bee: a common garden experiment. *Behavioral Ecology and Sociobiology*, *66*(5), 743-756.

Hepburn, H. R., Radloff, S. E., Otis, G. W., Fuchs, S., Verma, L. R., Ken, T., ... & Wongsiri, S. (2005). *Apis florea*: morphometrics, classification and biogeography. *Apidologie*, *36*(3), 359-376.

Hepburn, H.R., Pirk, C.W.W., Duangphakdee, O. (2014) Honey bee nests. composition, structure, function. Springer Verlag, Berlin, Heidelberg, New York

Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G., & Jarvis, A. (2004). The WorldClim interpolated global terrestrial climate surfaces. Version 1.3.

Meixner, M. D. (2010). A historical review of managed honey bee populations in Europe and the United States and the factors that may affect them. Journal of invertebrate pathology, 103, S80-S95.

Oleksa, A., & Tofilski, A. (2015). Wing geometric morphometrics and microsatellite analysis provide similar discrimination of honey bee subspecies. *Apidologie*, *46*(1), 49-60.

R Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Rinderer, T. E., Koeniger, N., Tingek, S., Mardan, M., & Koeniger, G. (1989). A morphological comparison of the cavity dwelling honeybees of Borneo *Apis koschevnikovi* (Buttel-Reepen, 1906) and *Apis cerana* (Fabricius, 1793). *Apidologie*, *20*(5), 405-411.

Rohlf, F. J. (2005). tpsDig, digitize landmarks and outlines, version 2.05. Department of Ecology and Evolution, State University of New York at Stony Brook.

Ruttner, F., Tassencourt, L., & Louveaux, J. (1978). Biometrical-statistical analysis of the geographic variability of *Apis mellifera* LI Material and methods. *Apidologie*, *9*(4), 363-381.

Ruttner, F. (1986). Geographical variability and classification. *Bee genetics and breeding*, 23-56.

Ruttner, F. (1988) Biogeography and taxonomy of honey bees. Springer-Verlag, Berlin

Seeley, T. D. (2014). *Honey bee ecology: a study of adaptation in social life*. Princeton University Press.

Tarpy, D. R., Delaney, D. A., & Seeley, T. D. (2015). Mating frequencies of honey bee queens (*Apis mellifera* L.) in a population of feral colonies in the northeastern United States. PLoS One, 10(3), e0118734.

Tofilski, A. (2008). Using geometric morphometrics and standard morphometry to discriminate three honey bee subspecies. *Apidologie*, *39*(5), 558-563.

Zhu, X., Zhou, S., Xu, X., Wang, J., Yu, Y., Yang, K. C., Luo, Q., Xu, Y., Wang, S., & Zhou, B. (2017). Morphological differentiation in Asian honey bee (*Apis cerana*) populations in the basin and highlands of southwestern China. Journal of Apicultural Research, 56(3), 203-209.